METHODS FOR HYPERTEXT REPORTING IN A RELATIONAL DATABASE MANAGEMENT SYSTEM

COPYRIGHT NOTICE

BACKGROUND OF THE INVENTION

The present invention relates generally to information processing environments and, more particularly, to modeling information in a data processing system, such as a Database Management System (DBMS).

Computers are a powerful tool for the acquisition and processing of information. Computerized databases can be regarded as a kind of electronic filing cabinet or repository for collecting computerized data files; they are particularly adept at processing vast amounts of information quickly. As such, these systems serve to maintain information in database files or tables and make that information available on demand. Of these systems, ones which are of particular interest to the present invention are Relational Database Management Systems (RDBMSs).

The concept of relational databases is perhaps best introduced by reviewing the problems surrounding traditional or non-relational systems. In a traditional database system, the task of retrieving information of interest (i.e., answering a "database query") is left to the user; that is, the user must give detailed instructions to the system on exactly how the desired result is to be obtained.

Consider the example of a simple query: "Who are the teachers of student John Smith?" In a traditional system, several explicit instructions are required before the query can be answered. One instruction, for instance, is typically to instruct the system to allocate sections in memory for data to be read from a storage disk. Another command may tell the system which disk files to open and read into the allocated memory for processing. Still other commands may specify particular search strategies, such as use of specific indexes, for speeding up the result of the query. And still even further commands may be needed for specifying explicit links between two or more files so that their data may be combined. Thus, instead of just telling the system "what" is desired (i.e., the desired data result as expressed in a query expression), one must specify internal procedures (i.e., the "how") for obtaining the data. Even for a simple query, such as that above, the task is complex, tedious, and error-prone.

From the user's perspective, such details—ones directed to the physical implementation—are completely irrelevant; the user is interested only in the result. Thus, the lack of separation of logical operations from the physical representation of the data (i.e., how it is internally stored and accessed by the system) in traditional systems burdens users with unnecessary complexity. Moreover, as traditional database products employ proprietary data access procedures, knowledge of one product is not necessarily helpful in use of another. And where database systems differ, their practitioners cannot effectively communicate with one another.

In 1970, Dr. E. F. Codd invented the "relational model", a prescription for how a DBMS should operate. The relational model provides a foundation for representing and manipulating data, that is, a way of looking at data. The model includes three basic components: structure, integrity, and manipulation. Each will be described in turn.

The first of these, structure, is how data should be presented to users. A database management system is defined as "relational" when it is able to support a relational view of data. This means that data which a user can access and the operators which the user can use to operate upon that data are themselves relational. Data are organized as relations in a mathematical sense, with operators existing to accept relations as input and produce relations as output. Relations are perhaps best interpreted by users as tables, composed of rows (tuples) and columns (attributes).

Ideally, data in a relational system is perceived by users as tables and nothing but tables. This precludes the user from seeing explicit connections or links between tables, or having to traverse between tables on the basis of such links. It also precludes user-visible indexes on fields and, in fact, precludes users from seeing anything that smacks of the physical storage implementation. Thus, tables are a logical abstraction of what is physically stored.

The integrity aspect, on the other hand, dictates that every relation (i.e., table) should have a unique, primary key to identify table entries or rows. The integrity of the data for the user is of course crucial. If accuracy and consistency of the data cannot be achieved, then the data may not be relied upon for decision-making purposes.

Data manipulation, the last component, may be thought of as cut-and-paste operators for tables. Data manipulation is of course the purpose for which databases exist in the first place. The superiority of manipulating tables relationally (i.e., as a whole, or sets of rows) is substantial. Users can combine data in various tables logically by matching values in common columns, without having to specify any internal details or the order in which tables are accessed; this provides users with a conceptual view of the database that is removed from the hardware level. Non-relational DBMSs, in contrast, require complex programming skills that form an inherently unreliable means to interact with databases.

The general construction and operation of a database management system is known in the art. See e.g., Date, C., *An Introduction to Database Systems,* Volume I and II, Addison Wesley, 1990; the disclosures of which are hereby incorporated by reference.

Today, relational systems are everywhere—commonly seen operating in corporate, government, academic settings, and other shared environments. A typical installation will employ one of the popular UNIX-based RDBMS running on a minicomputer. By submitting queries to the DBMS from a remote terminal (e.g., using a SQL "query editor"), users are often able to handle many of their own data processing needs directly. Thus, relational technology is not only just another way to build a database system, but it also offers a set of underlying principles that provide very direct practical benefits to the user.

A chief aim of the RDBMS is to provide company management with timely reports from which meaningful business decisions can be made. If back orders, say, at a given branch store are higher than at other branches, prompt attention can rectify the position, but only if the reporting system clearly indicates the possible anomaly to those capable of taking action.

Traditionally, the database system has produced printed daily, weekly or monthly order-status reports for branches